

J Exp Criminol (2013) 9:389–409
DOI 10.1007/s11292-013-9190-3

Long-term follow-up of a preschool experiment

Lawrence J. Schweinhart

Published online: 15 October 2013

© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract

Objectives This study was designed to provide experimental evidence of the effects of a preschool program on young children living in poverty. It began as a program evaluation but now, half a century later, serves as a test of the long-term effects and return on investment of high-quality preschool education for young children living in poverty.

Methods This study was conducted in the U.S., beginning in the 1960s, and has generated data on study participants from birth through 40, with new data now being collected at age 50. The study used random assignment procedures to assign 123 children to a preschool program and a control group who receive no preschool program.

Results Program participants surpassed non-participants in intellectual performance at school entry, school achievement throughout schooling, commitment to schooling, high school graduation rate, adult employment rate and earnings, reduced childhood antisocial behavior, and reduced adult crime and incarceration. The program's return on investment was at least seven times as great as its operating cost.

Conclusions While these powerful results have been found not only in this study but in several similar studies, they have not been found in studies of larger preschool programs, such as the Head Start Impact Study. This discrepancy suggests that differences between the two types of programs account for the better results found in studies such as this one. Among these differences are highly qualified teachers, a valid child development curriculum, extensive engagement of parents, and regular assessment of program implementation and children's development

Keywords Preschool · Experiment · Long-term follow-up · Longitudinal · Perry Preschool · Crime prevention · Return on investment · School readiness · School achievement · High school graduation · Adult earnings · Adult employment

Background

The modern preschool program research tradition began in the 1960s. Many of these studies have compared a group of children that experienced a preschool program with

L. J. Schweinhart (✉)

HighScope Educational Research Foundation, 600 North River Street, Ypsilanti, MI 48198-2898, USA
e-mail: lschweinhart@highscope.org

a similar group of children that did not experience a preschool program. Others have compared similar groups of children that experienced different kinds of preschool programs. In several of these studies, children were randomly assigned to the groups and thus defined as experiments. This article focuses on one such experiment.

The preschool research tradition flourished through the 1960s and 1970s, with additional studies of preschool and parent education programs. The principal investigators brought a dozen of the best of these studies together to form the Consortium for Longitudinal Studies (1983) in the late 1970s. Collaborating in data collection and analysis, the Consortium found robust evidence of preschool program effects on children's intellectual performance at school entry, reduced need for placements in special education, and reduced need for retention in grade. The longest-lasting of the studies found that a greater percentage of preschool program graduates became high school graduates.

Ramey and his colleagues at the University of North Carolina at Chapel Hill began the Carolina Abecedarian Study in 1972 (Campbell et al. 2002). Of 111 infants from poor families, they randomly assigned 57 to a special program group and 54 to a typical child care group that used the child care arrangements in homes and centers that were prevalent there in the 1970s. The special program was a full-day, full-year day care program for children that lasted the 5 years from birth to elementary school. This was the first study to find preschool program benefits on participants' intellectual performance and academic achievement throughout their schooling. The program group's mean IQ was the same as that of the no-program group at study entry, but significantly higher from ages 3 through 21; the same was true for achievement test scores at 15. By 15, fewer of the program group than the no-program group had repeated a grade or received special services, and by 21, more of them had graduated from high school or received a GED certificate and attended a 4-year college. Fewer of the program group participants than the no-program group became teen parents. However, unlike the HighScope Perry Preschool Study, the program and no-program groups did not differ significantly in arrests by 19 (Clarke and Campbell 1998), perhaps because of the few arrests by that age or perhaps because the Abecedarian curriculum did not focus on children's responsibility as much as the HighScope Perry Preschool curriculum did. Cost-benefit analysis indicates that, discounted at 3 % annually, the program yielded a return to society of \$3.78 per dollar invested (Massé and Barnett 2002).

Beginning in 1985, the Chicago Longitudinal Study, conducted by Arthur Reynolds and his colleagues examined the effects of the Chicago Child-Parent Centers (CPC) program offered by the nation's third largest public school district (Reynolds et al. 2001). This program was citywide, with 1,539 low-income children (93 % African American, 7 % Hispanic) enrolled in 25 schools, 989 who had been in the CPC program and 550 who had not. Families in this study went to their neighborhood schools, and children were not randomly assigned to groups. Preschool-program group members attended a part-day preschool program at ages 3 and 4, while the no-preschool-program group did not. The preschool-program group did significantly better than the no-preschool-program group in educational performance and social behavior, with lower rates of grade retention and special education placement, followed by a higher rate of high school completion and lower rates of school dropout and juvenile arrests. Analysis of the costs and benefits of the program indicates that, discounted at 3 % annually, the program yielded \$7.10 return per dollar invested (Reynolds et al. 2002).

The pattern of findings from this research tradition to date is that some model preschool programs have strong short-term effects, long-term effects, and strong return on investment, that is to say, they are highly effective (Schweinhart 2011). The evidence for typical preschool programs is mixed, however. The Head Start Impact Study (Puma et al. 2012) found that typical Head Start preschool programs have weak short-term effects with small likelihood of long-term effects and strong return on investment. However, several local and state preschool programs have been found to be highly effective in the short term (Barnett et al. 2005, 2013; Gormley et al. 2008; Schweinhart et al. 2012; Weiland and Yoshikawa 2013).

This article elaborates on a study that we have conducted that is central to drawing the conclusion that a high-quality preschool program for children of low-income families has long-term effects, including crime prevention—the HighScope Perry Preschool Study (Schweinhart et al. 2005). It presents how the study design warrants the findings, the array of both short-term and long-term findings, and how the short-term effects led to the long-term effects. Then it considers the program ingredients that led to these effects.

The study

To conduct this study, staff identified 123 young African American children in Ypsilanti, Michigan, USA, living in poverty and assessed to be at high risk of school failure. They used randomizing procedures to assign them to a program group and a no-program group, operated a high-quality preschool program for the program group at ages 3 and 4, and collected data on both groups annually from ages 3 through 11 and at ages 14, 15, 19, 27, and 40.

The study participants

This study has followed the lives of 123 persons who originally lived in the attendance area of the Ypsilanti, Michigan, school district's Perry Elementary School, a predominantly African American neighborhood in a low-income part of town. Project staff identified a pool of children for the study sample from a census of the families of students then attending Perry School, referrals by neighborhood groups, and door-to-door canvassing. They selected families of low socioeconomic status and children with low intellectual performance at study entry who showed no evidence of organic handicap. Only three families with children identified for the study refused to participate in it.

Although 128 children were originally selected for the study, four did not complete the preschool program because they moved away and one child died shortly after the study began, so that the longitudinal study had 123 participants. Children entered the study in five successive cohorts annually from the fall of 1962 to the fall of 1965. The relative homogeneity of the children in the sample makes it unlikely that annual selection had any effect on the randomization process. Staff randomly divided the children in each cohort into those enrolled in the preschool program and those not enrolled in any preschool program. Program-group children attended the preschool program at 3 and 4, except for the first class of children, who attended only at 4.

All study participants were African American, as was almost everyone in the Perry School neighborhood at the time they attended. Limiting the sample to African American children removed racial differences from the design, but raises a question about generalizing findings to other races, including the majority of poor children in the U.S., who are white. Obviously, poor African American children and poor white children differ in many important ways, but the program addressed intellectual, social, and physical abilities common to all races rather than racially specific characteristics; so it is reasonable to believe that it would have had similar effects on poor children of any race. Indeed, having positive effects on African American children may have been even harder to achieve, since regardless of preschool program participation, they and their families faced racial prejudice and discrimination in schooling, housing, and employment.

Program and no-program groups

The scientific strength of this study, its ability to assess preschool program effects even many years later, is due to an experimental design in which study participants were assigned by randomizing procedures to one of two groups: a "program group" enrolled in the preschool program or a "no-program group", which was not enrolled in any preschool program and received no special treatment other than data collection. After selecting children for a cohort in the study sample each fall from 1962 to 1965, project staff assigned them to program and no-program groups as follows. They identified pairs of study participants matched on initial Stanford-Binet intellectual performance scores (IQ, Terman and Merrill 1960) and assigned pair members to either of two undesignated groups. They exchanged one or two pair members per cohort to insure that the groups were matched on mean socioeconomic status, mean intellectual performance, and percentages of boys and girls. They randomly assigned one group to the program condition and the other to the no-program condition. In addition, as part of the initial assignment procedure in later cohorts, they exchanged several pair members in each cohort to reduce the number of children of employed mothers in the program group, because it was found difficult to arrange home visits with them. Also, they assigned younger siblings to the same group (program or no-program) as their older siblings, to prevent the preschool program from affecting siblings in the no-program group. Statistical analyses indicate that these exchanges did not appreciably affect the findings of this study (Heckman et al. 2010b). The initial assignments plus reassignments of children with employed mothers and younger siblings led to a program group of 58 children and a no-program group of 65 children.

The assignment procedures make it highly probable that group comparisons reflect the effects of the preschool program. Comparisons indicating that groups were not significantly different on various background characteristics make it even more likely. As shown in Table 1, the two groups differed significantly (with a two-tailed probability of less than 0.05) at study entry on only one background variable. The program-group members had significantly more employed mothers than the no-program-group members (9 vs. 31 %), because of the exchanges that moved them from the program group to the no-program group. However, program and no-program groups did not differ significantly in their percentages of mothers employed when study participants were 15 years old. Further, maternal employment at study entry

Table 1 Background variables at study entry, by preschool experience

Variable	Program group		No-program group		Effect size
	Mean/%	SD	Mean/%	SD	
Stanford-Binet IQ at study entry	79.6	5.9	78.5	6.9	0.16
Family socioeconomic status	9.0	1.2	8.6	1.1	0.36
Mother's years of schooling	9.5	2.4	9.4	2.0	0.05
Father's years of schooling ^a	8.4	2.3	8.8	2.5	-0.02
Children in family	4.9	2.5	4.8	2.8	0.06
Older siblings	2.8	2.4	3.0	2.8	0.08
Child's age in months at study entry	42.7	6.2	41.9	5.9	0.14
Mother's age in years at study entry	29.6	6.2	28.7	6.9	0.13
Rooms in home at study entry	5.2	1.2	5.2	1.6	0.01
Persons per room at study entry	1.2	0.3	1.4	0.6	-0.33
Participant's gender					
Male	57 %		60 %		0.06
Female	43 %		40 %		
Family configuration at study entry					
Two-parent	55 %		51 %		0.09
Single-mother	45 %		49 %		
Employment of parents					
Father and mother employed	5 %		9 %		0.44*
Father alone employed	44 %		35 %		
Single mother employed	4 %		22 %		
No parent employed	47 %		34 %		
Father's employment level ^b					
Managerial or skilled	3 %		2 %		0.29
Semiskilled	7 %		3 %		
Unskilled	36 %		40 %		

Table adapted from Schweinhart et al. 2005, pp. 30–31. Copyright 2005 by HighScope Educational Research Foundation. Program group $n=56-58$, no-program group $n=63-65$; data source is the initial parent interview; the p values are based on t tests of means and Chi-square statistics of percentages

* $p<0.05$, two-tailed

^a Program group $n=34$, no-program group $n=34$

^b All employed mothers reporting were at the unskilled level

was not correlated more than 0.13 with any of the major outcome variables. Nevertheless, it was used as a covariate in the outcome analyses reported here.

The preschool program

The program's teachers conducted daily 2½-hour classes for children on weekday mornings and made weekly 1½-hour home visits to each mother and child on weekday afternoons. The 30-week school year began in mid-October and ended in

May. Of the 58 children in the program group, the 13 in the first class participated in the program for one school year at 4 and the 45 in subsequent classes participated in the program for two school years at 3 and 4. Successive program group classes attended the program together, the older class at 4, the younger at 3. Thus, the four teachers in the program served 20 to 25 children each school year, forming a child–teacher ratio varying from 5.00 to 6.25 children per teacher. This ratio was set to accommodate the demands of the weekly home visits, that is, a visit to one or two families per weekday afternoon. Friday afternoons were generally used for staff training and project meetings. Between 1962 and 1967, ten teachers certified to teach in elementary, early childhood, and special education served in the program's four teaching positions. Over the years, seven white and three African American women served as teachers, and at least one African American teacher was always on the staff.

Program staff developed a systematic approach to classroom and home visit activities based on the idea that both teachers and children should have a major role in defining and initiating children's learning activities. Originally called the Cognitively Oriented Curriculum to distinguish it from approaches that did not include a systematic emphasis on cognitive development (Weikart et al. 1971), the education model was later named HighScope (Epstein and Hohmann 2012). From the beginning, the program staff explicitly sought to support the development of young children's cognitive and social skills through individualized teaching and learning. They continued to develop the educational model as the program operated from 1962 through 1967, building on insights from their classroom experience and their review of the studies of Jean Piaget and others (as summarized by Piaget and Inhelder 1969).

The HighScope early childhood educational model, used in the Perry Preschool classroom and home visits, was and is an open framework of educational ideas and practices based on the natural development of young children. Drawing on the child development ideas of Jean Piaget, Lev Vygotsky, and John Dewey, it emphasizes the idea that children are intentional learners, who learn best from activities that they themselves plan, carry out, and review afterwards. Adults introduce new ideas to children through adult-initiated small- and large-group activities. Adults observe, support, and extend the children's play as appropriate. Adults arrange interest areas in the learning environment—art area for art materials, block area for blocks, dress-up area for clothing, reading area for books, and so forth. They maintain a daily routine that permits children to plan, carry out, and review their own activities; and engage in small- and large-group times, cleanup time, snack time, and outside time. They join in children's activities, asking appropriate questions that extend their plans and help them think about their activities. They add complex language to the discussion to expand the child's vocabulary. Using key developmental indicators derived from child development theory as a framework, adults encourage children to make choices, solve problems, and engage in activities that contribute to their intellectual, social, and physical development.

Data collection

The HighScope Perry Preschool study has accumulated an unusually rich and comprehensive data set on young people growing up in poverty, with variables representing their status from birth through childhood and adolescence to early adulthood and midlife. The many variables encompass demographic characteristics,

test performance throughout childhood and adolescence, school success, crime, socioeconomic success, health, family, and personal development. The study itself provides some evidence of the validity of these measures in that they relate to each other as expected. Earlier tests and scales provide evidence of their internal consistency, but such information is not available for single items or official records of schooling, arrests, and sentences. The planned age 50 study will build on the age 40 study, expanding coverage of health and personality characteristics.

Attrition in the study sample has been very low. Across the 48 measures of 123 cases, a median seven cases (5.7 %) were missing. At the age 40 interview, of the 123 original study participants, 112 were interviewed, four living ones were not, and seven were deceased. Criminal justice and social services records were considered to have no missing cases because the names of all the study participants were included in these searches, and the lack of a record indicated no arrests or no social services, although it is possible of course that some information was recorded in agencies whose records were not searched. School records were found for all but 11 cases (8.9 %). The low rates of missing data mean that attrition had a negligible effect on either sample representativeness or group comparisons on outcome variables. At ages 14–15, when missing data ranged from 11 to 33 %, analyses of differential attrition found no effect on five background variables (gender, initial IQ, socioeconomic status, single parenthood, or mother's schooling (Schweinhart and Weikart 1980).

Methods of analysis

The analytic techniques presented in this report are based on comparisons of the program group and the no-program group with statistical adjustments to compensate for the effects of seven background covariates—participants' gender, Stanford-Binet IQ at study entry, mother's schooling, father's occupational status, household rooms per person, mother's employment, and father at home (i.e., single motherhood). Five of these variables had statistically significant relationships with one or more key outcome variables. Father at home was included because of its policy relevance and nearly significant relationship with monthly earnings at 40. Mother's employment was included because of its relationship with preschool experience induced by the group assignment procedure. Group differences in outcome variables, statistically adjusted by the seven covariates, were examined by binary logistic regression analysis for dichotomous variables, ordinal regression analysis for ordinal variables, and ordinary least-squares regression analysis for normally distributed interval variables, such as tested performance. Except for tested performance, the distributions of most outcome variables were not normal but L-shaped, that is to say, positively skewed to the right. We truncated these distributions by dividing the variable into equal-sized segments, and then analyzed them as above. Throughout this article, a group difference is identified as significant if it is statistically significant with a one-tailed probability of chance occurrence of less than 0.05, using one-tailed probabilities because the study's hypotheses are clearly directional at this point.

Heckman et al. (2010b) conducted a general reanalysis of the HighScope Perry Preschool Study, using innovative statistical procedures to correct for the study's small sample size, departures from random assignment, and multiple hypothesis testing. While, as might be expected, specific findings differed, the general result was to confirm the study's internal validity and provide greater scientific confidence in its results.

Findings

The variety of findings of program effects through 40 spans the domains of education, economic performance, crime prevention, and family and health (Schweinhart et al. 2005).

Education and economic status

The program group significantly outperformed the no-program group on various intellectual and language tests from their preschool years up to 7; school achievement tests at 7, 8, 9, and 14; and literacy tests at 19 and 27 (Schweinhart et al. 1993; Schweinhart and Weikart 1980; Schweinhart et al. 2005). The program group had significantly better attitudes towards school than the no-program group at 15 (seven items, $r_{\alpha}=0.634$) and 19 (16 items, $r_{\alpha}=0.799$). The program group significantly outperformed the no-program group on highest level of schooling completed (77 vs. 60 % graduating from high school or adult high school or obtaining a GED certificate). A much larger percentage of program than no-program females graduated from regular high school (88 vs. 46 %). This difference was related to earlier differences between program versus no-program females in the rates of treatment for mental impairment (8 vs. 36 %) and retention in grade (21 vs. 41 %).

Significantly more of the program group than the no-program group were employed at 27 (69 vs. 56 %) and 40 (76 vs. 62 %). Oddly, this advantage favored females at 27 (80 vs. 55 %) but males at 40 (70 vs. 55 %). The program group had significantly higher earnings than the no-program group, with medians of \$12,000 versus \$10,000 at 27 and \$20,800 versus \$15,300 at 40, as well as monthly at both ages.

Significantly more of the program group owned their own homes at 27 (27 vs. 5 %) and at 40 (37 % vs. 28 %). At 40, program males paid significantly more per month for their dwelling than did no-program males. Significantly more of the program group than the no-program group had a car at 40 (82 vs. 60 %), especially males (80 vs. 50 %), and at 27 (73 vs. 59 %). Indeed, at 27, a significantly larger percentage of the program group than the no-program group had a second car (30 vs. 13 %), especially males (36 vs. 15 %). At 40, significantly more of the program group than the no-program group had a savings account (78 vs. 50 %), especially males (73 vs. 36 %). At 27, significantly fewer of the program group than the no-program group reported receiving social services at some time in the previous 10 years (59 vs. 80 %).

Self- and teacher-reported misconduct

According to the ratings of kindergarten through third-grade teachers, the program group engaged in personal and school misconduct significantly less frequently than the no-program group at 6 through 9 ($p<0.05$, one-tailed). Personal misconduct (called personal behavior in some other reports of this study) had six items—absences or trancies, inappropriate personal appearance, lying or cheating, stealing, swearing or using obscene words, and poor personal hygiene ($r_{\alpha}=0.754$). School misconduct had 12 items, such as blaming others for trouble, being resistant to teacher, attempting to manipulate adults, and influencing others toward troublemaking ($r_{\alpha}=0.762$). All the items on both scales were scored very infrequently, infrequently, sometimes, frequently, or very frequently.

Program group members self-reported noticeably but not significantly fewer arrests than no-program group members, both up to 27 (24 vs. 35 %) and from 26 to 40 (30 vs. 46 %). The difference only reached statistical significance for program versus no-program males from 26 to 40 (33 vs. 60 %, $p < 0.05$, one-tailed). Comparing these statistics to the arrest statistics in Table 2, it can be seen that individuals under-reported whether they were ever arrested, so that the self-reported arrest group percentages were considerably less than the recorded arrest group percentages: 48 % of the program group and 57 % of the no-program group arrested up to 27 and 55 % of the program group and 71 % of the no-program group from 28 to 40. Similarly, program group members self-reported significantly fewer acts of misconduct than no-program group members by 15 (43 vs. 65 % reporting three or more such acts), but not significantly fewer at 19, 27, or 40.

Self- and teacher-reported misconduct over time complements this presentation of findings based on official criminal records. No one source of information on antisocial behavior is without challenges to its validity. Some see arrests as indicating more about the behavior of police towards various racial and ethnic groups than about the behavior of those arrested. In its simple form, this argument is tangential to the validity of the findings reported here because all study participants were African American. To apply, the argument would have to maintain that the program group engaged in behavior less likely to prejudice police than did the no-program group. Criminal behavior is the parsimonious explanation. On the other hand, it is obvious that many people commit crimes for which they are not arrested, and self-report is the reasonable way to count such crimes—if, and it is a big if—those committing such crimes report them accurately to the interviewer. However, social desirability encourages respondents to undercount crimes, and memory becomes less precise as the number of crimes exceeds two or three. In addition, the best self-reported indicator of crime is number of arrests; asking respondents to characterize actions for which they were not arrested as criminal or not requires them to make judgments for which they are neither legally competent nor particularly disposed to make.

Crime

The study presents strong evidence of a lifetime effect of the HighScope Perry Preschool program in preventing total arrests and arrests for violent, property, and drug crimes and subsequent prison or jail sentences.

Table 2 presents findings for arrests and general types of crimes cited at arrest by 40, further broken out by age. Compared to the no-program group, the program group had significantly fewer arrests by 40, specifically adult arrests by 27, but no fewer juvenile arrests or arrests from 28 to 40. The odds of lifetime arrests were 46 % lower for the program group than the no-program group—a little less than the odds reduction for juvenile arrests (despite the lack of a significant difference for this variable) and adult arrests through 27.

- 55 % of the no-program group but only 36 % of the program group were arrested five or more times in their lifetimes
- 29 % of the no-program group but only 7 % of the program group were arrested five or more times as adults by 27

Table 2 Arrests and crimes cited at arrest, by age by group

Arrests/crimes cited	Number	Lifetime by age 40			Juvenile by age 19			Adult by age 27			Adult ages 28–40		
		Program group	No-program group	Odds ratio	Program group	No-program group	Odds ratio	Program group	No-program group	Odds ratio	Program group	No-program group	Odds ratio
Arrests	0	29 %	17 %	0.54*	85 %	75 %	0.47	52 %	43 %	0.51*	45 %	29 %	0.83
	1–4	35 %	28 %		12 %	20 %		41 %	28 %		26 %	34 %	
	5–10	22 %	24 %		3 %	5 %		2 %	14 %		19 %	22 %	
	11–49	14 %	31 %		0 %	0 %		5 %	15 %		10 %	15 %	
Violent	0	67 %	52 %	0.46*	95 %	94 %	0.94	74 %	71 %	0.80	86 %	69 %	0.17**
	1–2	22 %	28 %					21 %	20 %		14 %	22 %	
	3–11	10 %	20 %					5 %	9 %		0 %	9 %	
	0	64 %	42 %	0.41**	88 %	81 %	0.60	72 %	65 %	0.59	85 %	68 %	0.39*
Property	1–2	26 %	32 %					23 %	15 %		8 %	21 %	
	3–20	10 %	26 %					5 %	20 %		7 %	11 %	
	0	86 %	66 %	0.38*	100 %	98 %	0.00	91 %	75 %	0.34*	90 %	77 %	0.44
	1–2	7 %	23 %					6 %	22 %		7 %	18 %	
Drug	3–9	7 %	11 %					3 %	3 %		3 %	5 %	

Table adapted from Schweinhart et al. 2005, p. 98. Copyright 2005 by HighScope Educational Research Foundation. Statistical tests are one-tailed. They and odds ratios are based on ordinal regression analysis, adjusted for the effects of participants' gender, Stanford-Binet IQ at study entry, mother's schooling, mother's employment, father at home, father's occupation status, and household rooms per person. The odds ratio compares the odds of the program group having more crimes (using the ordinal categories) to the no-program group having more crimes

* p (one-tailed) <0.05 ; ** p (one-tailed) <0.01

Compared to the no-program group, the program group had significantly fewer lifetime arrests for violent, property, and drug crimes by 40, but no fewer arrests for other crimes. The odds of arrests for violent, property, and drug crimes were 54–62 % lower in the program group than in the no-program group. Over their lifetimes to 40,

- 48 % of the no-program group but only 32 % of the program group were arrested for one or more violent crimes.
- 58 % of the no-program group but only 36 % of the program group were arrested for one or more property crimes.
- 34 % of the no-program group but only 14 % of the program group were arrested for one or more drug crimes.

Because, in general, males commit more crimes than females, we examined the crime outcome variables for group-by-gender interaction effects, that is, patterns in which a program effect was found for males but not females or females but not males. The regression analyses found no group-by-gender interaction effects for any of the crime or sentencing variables. Table 3 examines this question for arrests and general types of crime by 40, using the less stringent standard of whether statistically significant group differences were also statistically significant for males, females, or both taken separately. Recall that overall program and no-program group differences for all four of these variables were statistically significant.

Compared to no-program males, program males had significantly fewer arrests overall and significantly fewer arrests for property and drug crimes. Compared to no-program females, program females had significantly fewer arrests for violent and property crimes. Arrests for property crimes showed a significant difference among males and females separately. The biggest difference was for female arrests for violent crimes, 8 % for program females versus 27 % for no-program females. While the group difference in arrests for violent crimes by 40 was not significant for males, it was significant for males from 28 to 40, with 21 % of program males with one such arrest versus 43 % of no-program males with one or more such arrest (odds ratio=0.14, $p<0.01$).

Table 4 presents group comparisons on adult criminal convictions and sentences, by 40 and broken out up to 27 and at 28–40. Program group members were convicted and sentenced to significantly fewer months in prison or jail by 40 than were no-program group members. The odds of the program group spending time in prison or jail were 52 % less than they were for the no-program group. Over their lifetimes, 52 % of the no-program group but only 28 % of the program group were sentenced to any time in prison or jail.

The program group had less sentencing than the no-program group on every measure of sentencing, but not to a statistically significant extent for any other single measure—undropped misdemeanor cases, convicted felony crimes, sentenced to prison for felonies, months sentenced to probation, or months served in prison.

Table 4 also presents group comparisons on criminal sentences through 27 and from 28 to 40. Compared to the no-program group, the program group had no significant differences in sentencing by 27. Compared to the no-program group, the program group had significantly fewer members sentenced to prison for felonies from 28 to 40, was sentenced to significantly fewer months in prison or jail, and served significantly fewer months in prison, but there were no significant group differences in undropped misdemeanor cases, convicted felony crimes, or months sentenced to

Table 3 Arrests and crimes cited at arrest by age 40, by gender by group

Crime cited at arrest	Program males	No-program males	Odds ratio	Program females	No-program females	Odds ratio
Arrests						
0	18 %	5 %	0.45*	44 %	35 %	0.52
1–4	37 %	26 %		32 %	31 %	
5–10	21 %	28 %		24 %	19 %	
11–49	24 %	41 %		0 %	15 %	
Violent crimes						
0	49 %	38 %	0.47	92 %	73 %	0.12*
1–2	33 %	31 %		8 %	23 %	
3–11	18 %	31 %		0 %	4 %	
Property crimes						
0	52 %	28 %	0.43*	80 %	62 %	0.20*
1–2	30 %	36 %		20 %	27 %	
3–20	18 %	36 %		0 %	11 %	
Drug crimes						
0	82 %	51 %	0.34*	92*	89 %	0.78
1–2	6 %	39 %		8 %	0 %	
3–9	12 %	10 %		0 %	11 %	

Table adapted from Schweinhart et al. 2005, p. 89. Copyright 2005 by HighScope Educational Research Foundation. Statistical tests are one-tailed. Statistical tests and odds ratios (but not group percentages) are based on ordinal regression analysis, adjusted for the effects of participants' gender, Stanford-Binet IQ at study entry, mother's schooling, mother's employment, father at home, father's occupation status, and household rooms per person. The odds ratio compares the odds of the program group having more crimes (using the ordinal categories) to the no-program group having more crimes

probation. Compared to the odds for the no-program group, the odds of the program group being sentenced to prison for felonies were 78 % less, of being sentenced to prison or jail were 59 % less, and of serving months in prison were 63 % less.

- 25 % of the no-program group but only 7 % of the program group were sentenced to prison for felonies from 28 to 40.
- 43 % of the no-program group but only 19 % of the program group were sentenced to prison or jail from 28 to 40.
- 21 % of the no-program group but only 9 % of the program group served time in prison from 28 to 40.

Crime patterns

Several questions can be raised about the evidence presented herein. The first is whether it truly leads to the conclusion that this program prevented crime. The second is what such a conclusion really means. Table 2 shows that the evidence is strong, but not totally consistent with this conclusion. It is always possible to focus on variation rather than central tendency. While significant group differences were found for lifetime arrests and

Table 4 Adult criminal sentences, by age by group

Variable	By age 40			By age 27			Ages 28–40		
	Program group	No-program group	Odds ratio	Program group	No-program group	Odds ratio	Program group	No-program group	Odds ratio
Undropped misdemeanors									
0	38 %	29 %	0.61	57 %	46 %	0.52	54 %	42 %	0.83
1–4	34 %	25 %		36 %	31 %		24 %	31 %	
5–9	19 %	25 %		5 %	17 %		19 %	18 %	
10–27	9 %	21 %		2 %	6 %		3 %	9 %	
Convicted felony crimes									
0	74 %	65 %	0.70	79 %	75 %	0.98	83 %	74 %	0.57
1–2	14 %	17 %		14 %	14 %		15 %	17 %	
3–13	12 %	18 %		7 %	11 %		2 %	9 %	
Prison sentence for felonies	19 %	31 %	0.54	19 %	20 %	1.20	7 %	25 %	0.22**
Months sentenced to prison or jail									
0	72 %	48 %	0.48*	76 %	72 %	1.13	81 %	57 %	0.41*
1–24	16 %	29 %		12 %	17 %		14 %	29 %	
25–510	12 %	23 %		12 %	11 %		5 %	14 %	
Months sentenced to probation									
0	74 %	63 %	0.64	88 %	81 %	0.73	81 %	74 %	0.69
1–24	16 %	19 %		5 %	11 %		14 %	18 %	
25–144	10 %	18 %		7 %	8 %		5 %	8 %	
Prison months served									
0	86 %	78 %	0.62	86 %	84 %	0.92	91 %	79 %	0.37*
1–100	7 %	8 %		12 %	8 %		2 %	12 %	
101–250	7 %	14 %		2 %	8 %		7 %	9 %	

Table adapted from Schweinhart et al. 2005, p. 99. Copyright 2005 by HighScope Educational Research Foundation. Statistical tests are one-tailed. They and odds ratios are based on ordinal regression analysis, adjusted for the effects of participants' gender, Stanford-Binet IQ at study entry, mother's schooling, mother's employment, father at home, father's occupation status, and household rooms per person. The odds ratio compares the odds of the program group having more crimes (using the ordinal categories) to the no-program group having more crimes

types of crimes, they were not found for juveniles and varied for adults to 27 and at 28 to 40. The percentage-point arrest advantage of the program group over the no-program group varied from 9 to 16 points between 19 and 40. Similarly, the percentage-point advantage of the program group over the no-program group on self-reported misconduct was 22 points at 15, but nine points or less at 19, 27, and 40. However, this emphasis on differences and statistical significance misses the broader pattern: every single odds ratio of arrests, crimes, or sentences by 40 favored the program group over the no-program group. The same was true of all but three of the odds ratios of arrests, crimes, and sentences by 27 and from 28 to 40. In these three instances (other crimes from 28 to 40, sentenced to prison for felonies by 27, and months sentenced to prison or jail by 27), the

unadjusted percentages were lower for the program group than for the no-program group. The consistency of this pattern is strong evidence of a crime prevention effect.

With respect to the meaning of this conclusion, Table 2 suggests that the preschool program's crime prevention effect centered on violent, property, and drug crimes. Regarding specific types of crimes, the program effect was strongest for assault and/or battery, larceny under \$100, dangerous drugs, and disorderly conduct or disturbing the peace. These types of crimes signal a lack of impulse control. Use of dangerous drugs also indicates a serious disregard for long-term consequences. With its daily routine of children planning, doing, and reviewing their activities, the preschool program focused on strengthening their abilities to make decisions and plan their lives intelligently. Many violent, property, and drug crimes result from bad decision-making, disregard for consequences, and a lack of impulse control. It would seem that the preschool program helped children develop these traits to a greater extent than they would have otherwise.

For this explanation to apply, impulse control must be a behavioral trait that can be influenced by preschool experience and then remain stable until the onset of opportunities for criminal activity. The preschool curriculum-based explanation above focuses on variables that are proximal to the antisocial behavior that is antecedent to crime, the type of variables often featured in crime theorizing. As such, it could apply both in the preschool setting and in their families of origin, influenced by home visits. Family and preschool setting may be seen as mutually reinforcing pathways for the behavioral complex in which people's social and antisocial behavior is embedded. Then the question is not which one is responsible for the development of antisocial behavior, but rather how much each of them contributes. Parents' involvement in home visits and subsequent childrearing places family as a potential mediator of the program effects. The involvement of some of the study participants in the preschool program makes it a potential fountainhead for recurrent cycles of success, motivation for success, and avoidance of misconduct portrayed in the causal model.

Why were the violent and property crime prevention effects somewhat stronger at 28 to 40 than up through 27? Close examination of the statistics shows that a few more preschool participants than non-participants stopped engaging in personal and property violence after 28. Perhaps the impulse control they learned by preschool participation combined with their middle adult life circumstances to reduce such crime.

Causal model

We developed a causal model that takes into account the temporal ordering of the variables (Schweinhart et al. 2005). The relatively small sample size limited the number of terms to eight or so and renders the model suggestive rather than definitive. The structural equation model was estimated with the AMOS program (version 4.1, Arbuckle 1999). It traces the significant relationships between pairs of variables and does not account for the variance attributable to the preschool program itself other than the direct effect of preschool experience on postprogram IQ.

As shown in Fig. 1, the model suggests that the following causal path might be the route through which preschool program effects are transmitted to arrests by 40.

- (1) Preschool experience directly improves study participants' early childhood intellectual performance, which is also positively predicted by family socioeconomic status.

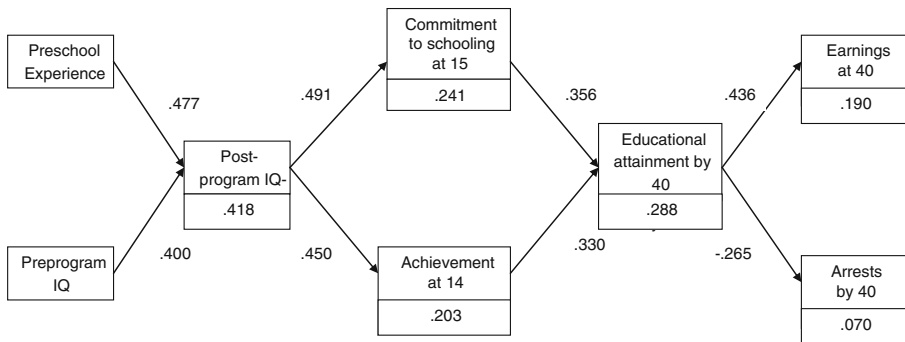


Fig. 1 A model of the paths from preschool experience to success at 40. Figure adapted from Schweinhart et al. 2005, p. 164. Copyright 2005 by HighScope Educational Research Foundation. Path coefficients are standardized regression weights that are statistically significant at $p < 0.01$; coefficients in each box are squared multiple correlations

- (2) Their early childhood intellectual performance improves their school motivation in elementary school.
- (3) Their early childhood intellectual performance and school motivation in elementary school reduce the number of years they spend in programs for children with mental impairment.
- (4) Because of study participants' early childhood intellectual performance and years spent in programs for children with mental impairment, they have higher literacy scores as they leave high school.
- (5) Study participants' school motivation leads them to complete a higher level of schooling.
- (6) Because they have completed a higher level of schooling, they have higher monthly earnings at 40 and fewer lifetime arrests.

Heckman et al. (2013) conducted another analysis of childhood mediators of adult effects in the study. They conducted separate analyses for males and females and did not focus on the preschool intellectual boost at its peak at 5, preferring to wait until it settled down at 7 to 9. Their analysis found that the preschool program improved children's intellectual performance, which later led to their improved school achievement and reduced use of welfare assistance by females. It decreased males' misconduct, thereby reducing their crime as adults. It improved females' social relationships while reducing their misconduct, improving their adult lives in various ways. They concluded that while intellectual development remains a goal of a high-quality preschool program, so are improving social relationships and reducing misconduct, especially for boys.

These mediator analyses do indicate that cognitive and school achievement are part of the explanation of preschool crime prevention, but so are various personality factors, such as social relationships and misconduct. The HighScope Curriculum itself suggests additional mediators such as curiosity, critical thinking, independent decision-making, and responsibility. Current thinking in this domain has expanded to executive function, impulse control, and anticipation of consequences. Preschool improvement in these traits may be inferred from

the existing data in this study, but could be more directly measured in future research.

Cost–benefit analysis

In constant 2013 dollars discounted at 3 %, the return to society was \$341,732 per participant on an investment of \$20,019 per participant (\$11,273 per participant per year) – \$16.14 per dollar invested (Belfield et al. 2006). Of that return, 80 % went to the general public – \$12.90 per dollar invested, and 20 % went to each participant – \$3.24 per dollar invested. Of the public return, 88 % came from crime savings, and the rest came from education and welfare savings and increased taxes due to higher earnings. A full 93 % of the public return was due to the large program effect of reduced crime rates for program males. Male program participants cost the public 41 % less in crime costs per person, \$967,420 less in undiscounted 2013 dollars over their lifetimes. Preschool program participants earned 14 % more per person than those who did not attend the preschool program – \$206,567 more over their lifetimes in undiscounted 2013 dollars.

Of particular interest in this article is the calculation of crime costs. Reductions in crime produce savings in victims' costs; criminal justice costs for policing, arrest, and sentencing; and incarceration and probation costs. These costs vary according to the type of crimes (e.g., murder, burglary) and the seriousness of the crime (felony or misdemeanor). Multiplying the incidence of each crime by its unit cost yields the total burden of crime. The crime incidents identified in this study were used to estimate numbers of lifetime crimes, including predictions of crime beyond 40 based on national data on arrest rate frequencies by age, even though criminal activity up to 40 represents 73–92 % of total lifetime criminal activity, with percentage depending on type of crime (Federal Bureau of Investigation 2002).

Many crimes do not result in an arrest. In the U.S. in 2002, 5.34 million violent crimes were reported by victims, but led to only 0.62 million arrests (Bureau of Justice Statistics 2002; Federal Bureau of Investigation 2002). We used these data to estimate that there were 3–14 crimes per arrest, depending on the type. We then estimated the costs of each crime to the victim and to the criminal justice system. Victim costs include expenses for medical treatments and to replace property or assets (even with insurance claims); lost productivity at work and at home; and reduced quality of life from pain, fear, and suffering. Estimates for these costs were derived from Miller et al. (1996). Criminal justice system costs for arrests, trials, and sentencing were adapted from Cohen (1998) and Cohen et al. (2004). Policing costs were estimated per crime, while sentencing and trial costs were estimated per arrest.

Heckman et al. (2010a) conducted a reanalysis of the costs and benefits of the Perry Preschool program that examines dozens of assumptions and produces hundreds of estimates of return on investment. It confirms and adds scientific confidence to the study's basic economic finding that the preschool program's economic benefits to taxpayers and program participants far exceed the cost of the program. Like the general reanalysis, the cost–benefit one builds on statistical procedures that correct for the study's small sample size and departures from random assignment. Unlike earlier analyses, it presents standard errors of the statistics and systematic sensitivity analysis and includes the deadweight costs of taxation. It finds a social rate of return of 14.3 % and a benefit–cost ratio of 7.1 to 1. These estimates are smaller than our

most recent estimate, about the same as earlier estimates, statistically significantly different from 0 % for both males and females, and above the historical rate of return on the U.S. stock market.

Discussion

This study initially detected effects on children's intellectual and language performance that suggested longer-term effects. In the elementary school years, it found that these effects did not persist beyond first grade. In the high school and adult years, it found that despite the disappearance of initial effects, the program had important long-term effects on schooling, economic performance, and crime.

Following this study, we conducted a curriculum comparison study to see if the effects of high-quality preschool education found for the HighScope preschool curriculum would also be found for a Direct Instruction or traditional Nursery School preschool curriculum (Schweinhart and Weikart 1997a, b). It found that their educational effects through 10 were quite similar, with an edge to the children who had a Direct Instruction preschool program. However, in the high school and early adult years, it found that the Direct Instruction program's educational advantage did not persist, and the HighScope and Nursery School programs had a host of emergent advantages on social outcomes, including better prevention of adult felonies and property crimes, emotional impairment or disturbance, and increased voluntarism. This study shows that long-term research can render a radically different assessment concerning whether or not programs are successful.

From these studies and the rest of the body of evidence on preschool programs, we conclude that model preschool programs and some local and state programs have the program ingredients needed to be highly effective, while the large, federal Head Start program does not currently have these program ingredients to the extent needed to be highly effective. We have identified these program ingredients as fully qualified or well-supervised teachers using a proven curriculum model, engaging parents as partners, and regularly assessing program implementation and children's development (Schweinhart 2011). Regular assessment of children's development provides critical feedback to program implementers regarding whether the program is on track towards providing children with the outcomes that will lead to long-term success—social and emotional as well as cognitive outcomes.

Long-term follow-up research enlarges beyond the perspective of short-term, survey research. Short-term research typically tells whether a program has its intended result. Long-term research typically examines a wider range of potential results and so offers a more nuanced and complex perspective of program results. A program can have an immediate effect but not a long-term one, or it can have no immediate effect but a long-term one nonetheless, or the situation can be more complex than that.

In this and similar studies, the preschool programs had an immediate effect on children's intellectual performance that lasted only a few years, leading to the notorious idea of a fadeout in effect on intellectual performance. Psychological traits and abilities, such as intelligence, are assumed to persist over time, just as physical objects persist over time. However, our senses directly perceive physical objects, whereas our senses do not directly perceive psychological traits and abilities, and

their existence must be inferred from the evidence of our senses. Test items may be selected for their stability over time under ordinary circumstances, but that does not mean that they are stable under unusual circumstances.

However, these studies found important long-term effects on high school graduation, adult earnings, and employment, and reduced crime. This pattern required a reconceptualization of the pathway to long-term effects. Clearly, the data did not support the original hypothesis of improved intellectual performance alone as the life-changer that would lead to these long-term effects. Participants were not experiencing better adult outcomes because they were smarter, at least in the way measured by intellectual tests.

The puzzle remains in this study. Our causal analysis identifies the temporary preschool program effect on intellectual performance as the gateway to longer-term effects. The analysis by Heckman et al. (2013) assigns a more limited role to intellectual performance as a path to school achievement and attaches more importance to social relations and personal conduct. Either way, these measures taken in the 1970s do not fully capture some key intended outcomes of the HighScope Curriculum, such as initiative, responsibility, and independent thinking, nor do they fully capture newly hypothesized mediators of long-term effects, such as executive function, self-regulation, and tenacity. The inability to reach back in time with emergent ideas is an inevitable frustration of long-term research.

The advantage of an original study design featuring a valid comparison of groups is that it can last for decades. This study has such a design. It was originally intended to be an evaluation lasting through the end of the preschool program. With the concern about fadeout, the question of enduring effects arose, propelling the study to 15. The discovery of lasting effects at 15 motivated further study at 19; effects at 19 prompted further study at 27; and effects at 27 prompted further study at 40 and, now, will prompt further study at 50. While this strategy has a sort of economy about it, it lacks the definition of a fully proactive longitudinal study.

In the preschool years, the focus was on children's intellectual and language tests. In the elementary school years, it shifted to school achievement tests, teacher ratings of children's social skills and motivation, and school placements. In the high school years, we added interviews of children and their parents and began the shift from educational psychology to social psychology and sociology. In adulthood, we focused on interviews and official records. We shifted from test performance to life performance—including high school graduation, earnings, employment, arrest records, and prison records.

Conclusions

Because of the random assignment and low attrition, this study has strong internal validity. The external validity or generalizability of the study findings extends to those programs that are reasonably similar to this program. A reasonably similar program is a preschool education program run by teachers with bachelors' degrees and certification in education, each serving up to eight children living in low-income families. The program runs two school years at 3 and 4 years of age, uses the HighScope

Curriculum, with daily classes of 2½ hours or more and teachers visiting families at least every 2 weeks. The curriculum comparison study (Schweinhart and Weikart 1997a, b) that followed this one suggest that curriculum had a lot to do with the findings.

The conclusion from the research presented in this article is that high-quality preschool programs for young children living in poverty contribute to their intellectual and social development in childhood and their school success, economic performance, and reduced commission of crime in adulthood. This study confirms that the long-term effects are lifetime effects.

The simple implication of this study is that all young children living in low-income families should have access to preschool programs that have features reasonably similar to the program used in this study. This study and others reviewed herein have motivated policymakers to invest in preschool programs. But because policymakers practice the art of compromise, these programs have seldom met the reasonable similarity standard. This study is a symbol of what government programs can achieve and inspires the passionate belief of those who want to believe in what government can accomplish and the passionate disbelief of those who want to believe that government cannot accomplish much of anything. But ultimately the public is not served by beliefs that are either too optimistic or too pessimistic. While the study can serve as grounds for debate, it is better to see it as a challenge. It shows what can be done, and the challenge is to do it. This study, the Abecedarian study, and the Chicago study lay down the same challenge: to do what we know how to do to prevent poverty from being a malevolent birthright handed down from generation to generation by the very schooling established to overcome it. The challenge is to provide high-quality preschool programs that include and actively engage low-income children so that these children get a fair chance to achieve their full potential to contribute to society.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Arbuckle, J. L. (1999). *AMOS*. Chicago: SmallWaters Corp. Now available from SPSS.
- Barnett, W. S., Jung, K., Youn, M.-J., & Frede, E. C. (2013). *Abbott preschool program longitudinal effects study: fifth grade follow-up*. New Brunswick: National Institute for Early Education Research, Rutgers University. Available online at <http://nieer.org/sites/nieer/files/APPLES%205th%20Grade.pdf>.
- Barnett, W. S., Lamy, C., & Jung, K. (2005). *The effects of state prekindergarten programs on young children's school readiness in five states*. New Brunswick: National Institute for Early Education Research, Rutgers University. Available online at nieer.org/docs/index.php?DocID=129.
- Belfield, C. R., Nores, M., Barnett, S. W., & Schweinhart, L. J. (2006). The HighScope Perry preschool program: cost-benefit analysis using data from the age-40 follow-up. *The Journal of Human Resources*, 41(1), 162–190.
- Bureau of Justice Statistics. (2002). *Criminal victimization in the United States*. Washington: U.S. Department of Justice.
- Campbell, F. A., Ramey, C. T., Pungello, E. P., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: young adult outcomes from the abecedarian project. *Applied Developmental Science*, 6, 42–57.
- Clarke, S. H., & Campbell, F. A. (1998). Can intervention early prevent crime later? The Abecedarian Project compared with other programs. *Early Childhood Research Quarterly*, 13, 319–343.

- Cohen, M. A. (1998). The monetary value of saving a high-risk youth. *Journal of Quantitative Criminology*, 14, 319–343.
- Cohen, M. A., Rust, R. T., Steen, S., & Tidd, S. T. (2004). Willingness-to-pay for crime control programs. *Criminology*, 42, 89–109.
- Consortium for Longitudinal Studies. (1983). *As the twig is bent...lasting effects of preschool programs*. Hillsdale: Erlbaum.
- Epstein, A. S., & Hohmann, M. (2012). *The HighScope preschool curriculum*. Ypsilanti: HighScope Press.
- Federal Bureau of Investigation. (2002). *Crime in the United States*. Washington: Department of Justice. Available online at <http://www.fbi.gov/ucr/ucr/htm#cuis>
- Gormley, W. T., Jr., Phillips, D. A., & Gayer, T. (2008). Preschool programs can boost school readiness. *Science*, 320(5884), 1723–1724.
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010a). The rate of the return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94, 114–128. Available online at www.ncbi.nlm.nih.gov/pmc/articles/PMC3145373/.
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. Q. (2010b). Analyzing social experiments as implemented: a reexamination of evidence from the HighScope Perry preschool program. *Quantitative Economics*, 1(1), 1–46. Available online at www.qeconomics.org/ojs/index.php/qe/article/view/2.
- Heckman, J. J., Malofeeva, L., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*. Available online at http://www.aeaweb.org/forthcoming/output/accepted_AER.php.
- Massé, L. N., & Barnett, W. S. (2002). A benefit-cost analysis of the Abecedarian early childhood intervention. In Levin, H., & McEwan P. (Eds.), *Cost effectiveness analysis in education: Methods, findings and potential*. 2002 Yearbook of the American Education Finance Association. Available online at <http://nieer.org/resources/research/AbecedarianStudy.pdf>.
- Miller, T. R., Cohen, M. A., & Wiersma, B. (1996). *Victim costs and consequences: a new look (National Institute of Justice Research Report)*. Washington: National Institute of Justice. (NCJ-155282).
- Piaget, J., & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
- Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., & Downer, J. (2012). *Third grade follow-up to the Head Start Impact Study final report, OPRE Report # 2012–45*. Washington: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Available online at http://www.acf.hhs.gov/sites/default/files/opre/head_start_report.pdf.
- Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest: A 15-year follow-up of low-income children in public schools. *Journal of the American Medical Association*, 285, 2339–2346.
- Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2002). Age 21 cost-benefit analysis of the Title I Chicago child–parent centers. *Educational Evaluation and Policy Analysis*, 4, 267–303. Version available online at <http://www.ssc.wisc.edu/irp/pubs/dp124502.pdf>.
- Schweinhart, L. J. (2011). How to make early childhood programs highly effective. In E. Zigler, W. Gilliam, & W. S. Barnett (Eds.), *The Pre-K debates: current controversies and issues*. Baltimore: Paul H. Brookes.
- Schweinhart, L. J., Barnes, H. V., & Weikart, D. P. (1993). *Significant benefits: the HighScope Perry preschool study through age 27*. Ypsilanti: HighScope Press.
- Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: the HighScope Perry preschool study through age 40*. Ypsilanti: HighScope Press. Summary available online at http://highscope.org/file/Research/PerryProject/specialsummary_rev2011_02_2.pdf.
- Schweinhart, L. J., Xiang, Z., Daniel-Echols, M., Browning, K., & Wakabayashi, T. (2012). *Michigan great start readiness program evaluation 2012: high school graduation and grade retention findings*. Ypsilanti: HighScope Educational Research Foundation. Available online at http://www.highscope.org/file/Research/state_preschool/MGSRP%20Report%202012.pdf.
- Schweinhart, L. J., & Weikart, D. P. (1980). *Young children grow up: effects of the Perry preschool program on youths through age 15. (Monographs of the HighScope educational research foundation, 7)*. Ypsilanti: HighScope Press.
- Schweinhart, L. J., & Weikart, D. P. (1997a). *Lasting differences: The HighScope Preschool Curriculum Comparison Study through age 23. (Monographs of the HighScope Educational Research Foundation, 12)*. Ypsilanti: HighScope Press.
- Schweinhart, L. J., & Weikart, D. P. (1997b). The HighScope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly*, 12, 117–143.

- Terman, L. M., & Merrill, M. A. (1960). *Stanford-Binet Intelligence Scale Form L-M: Manual for the third revision*. Boston: Houghton-Mifflin.
- Weikart, D. P., Rogers, L., Adcock, C., & McClelland, D. (1971). *The cognitively oriented curriculum: a framework for preschool teachers*. Urbana: University of Illinois.
- Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, 84 (6).

Lawrence Schweinhart is an early childhood program researcher and speaker throughout the United States and in other countries. He has conducted research at the HighScope Educational Research Foundation in Ypsilanti, Michigan, since 1975 and served as its president from 2003, becoming president emeritus in 2013. He has directed the HighScope Perry Preschool Study through age 40, the Michigan School Readiness Program Evaluation, HighScope's Head Start Quality Research Center, and the development and validation of the Child Observation Record. He received his Ph.D. in Education from Indiana University in 1975. He and his wife have two children and five grandchildren.